(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 048 970**
A2

(12) EUROPEAN PATENT APPLICATION

(21) Application number: 81107621.5

(22) Date of filing: 24.09.81

(51) Int. Cl.³: C 12 N 15/00, C 12 P 21/02, C 12 N 1/20, C 07 H 21/04, C 07 C 103/52, A 61 K 45/02 // C12R1/19, C12R1/125, C12R1/865

(30) Priority: 25.09.80 US 190799
11.08.81 US 291892

(43) Date of publication of application: 07.04.82 Bulletin 82/14

(84) Designated Contracting States: AT BE CH DE FR GB IT LI LU NL SE

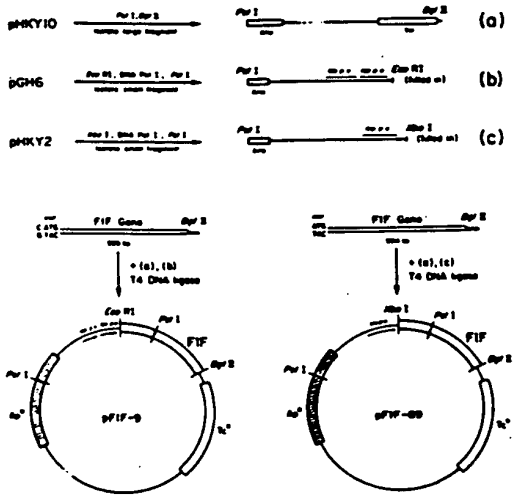
(71) Applicant: GENENTECH, INC., 460 Point San Bruno Boulevard, So. San Francisco California 94080 (US)

(72) Inventor: Crea, Roberto, 1520 Howard Avenue, Burlingame California (US)
Inventor: Goeddel, David Van Norman, 1449 Benito Avenue, Burlingame California (US)

(74) Representative: Lederer, Franz, Dr. et al, Patentanwälte Dr. Lederer Franz Meyer-Roxlau Reiner F. Lucile-Grahn-Strasse 22, D-8000 München 80 (DE)

(54) Polypeptides, process for their microbial production, intermediates therefor and compositions containing them.

(57) Microbially produced mature human fibroplast interferon and means used in its production, i.e. recombinant DNA molecules coding for the amino acid sequences of fibroplast interferon, vectors capable of expressing fibroblast interferon in microbial host organisms and host organisms transformed with these vectors.



EP 0 048 970 A2

Polypeptides, process for their microbial production, intermediates therefor and compositions containing them

The present invention relates to the field of recombinant DNA technology, i.e. to processes used in recombinant DNA technology and to products obtained by these processes.

In a more detailed aspect the present invention relates to polypeptides, specifically to mature human fibroblast interferon, to pharmaceutical compositions containing them and to a process for their preparation which comprises causing a culture of a microorganism transformed with a replicable microbial expression vehicle capable of expressing said polypeptides to grow up and express said polypeptides. The present invention also comprises the expression vehicles used in this process and the novel microorganisms containing these expression vehicles as well as the processes for their preparation. Finally, the invention relates to DNA sequences comprising sequences coding for the amino acid sequence of a mature human fibroblast interferon.

Background of the invention

Human fibroblast interferon (FIF) is a protein which exhibits antiviral as well as a wide range of other biological activities (for review see W.E. Stewart II, The Interferon System, Springer-Verlag, New York-Wien, 1979).

It has reportedly been purified to homogeneity as a single polypeptide with a molecular weight of 19000 - 20000 having a specific activity of 2-10 x 10^8 units/mg (E. Knight, Proc. Natl. Acad. Sci. USA 73, 520-523 [1976]; W. Berthold et al., J. Biol. Chem. 253, 5206-5212 [1978]). The sequence of the 13 NH_2-terminal amino acids of FIF has been determined to be Met-Ser-Tyr-Asn-Leu-Leu-Gly-Phe-Leu-Gln-Arg-Ser-Ser- (E.Knight et al., Science 207, 525-526 [1980]). Houghton et al. (Nucleic Acids Res. 8, 1913-1931 [1980]) have used synthetic deoxyoligonucleotides (predicted from this amino acid sequence) to determine the sequence of the 276 5'-terminal nucleotides of FIF mRNA. Taniguchi et al. (Nature 285, 547-549 [1980]; Gene 10, 11-15 [1980]) and Derynck et al. (Nature 285, 542-547 [1980]) have recently been able to identify the nucleotide sequence of cloned cDNA copies of FIF mRNA in E. coli and have deduced therefrom the complete amino acid sequence of human FIF including a 21 amino acids signal sequence. The mature peptide is 166 amino acids long. Finally, Taniguchi et al. (Proc. Natl. Acad. Sci. USA 77, 5230-5233 [1980]) have constructed a plasmid that directs expression in E. coli of the human FIF gene yielding mature FIF.

With the advent of recombinant DNA technology, the controlled microbial production of an enormous variety of useful polypeptides has become possible. Already in hand are bacteria modified by this technology to permit the production of such polypeptide products such as somatostatin, the A and B chains of human insulin, human growth hormone (Itakura et al., Science 198, 1056-1063 [1977]; Goeddel et al., Nature 281, 544-548 [1979]). More recently, recombinant DNA techniques have been used to occasion the bacterial production of proinsulin, thymosin α_1 and leukocyte interferon.

The workhorse of recombinant DNA technology is the plasmid, a non-chromosomal loop of double-stranded DNA found in bacteria and other microbes, oftentimes in multiple

copies per cell. Included in the information encoded in the plasmid DNA is that required to reproduce the plasmid in daughter cells (i.e., a "replicon") and ordinarily, one or more selection characteristics such as, in the case of bacteria, resistance to antibiotics which permit clones of the host cell containing the plasmid of interest to be recognized and preferentially grown in selective media. The utility of plasmids lies in the fact that they can be specifically cleaved by one or another restriction endonuclease or "restriction enzyme", each of which recognizes a different site on the plasmidic DNA. Thereafter heterologous genes or gene fragments may be inserted into the plasmid by endwise joining at the cleavage site or at reconstructed ends adjacent to the cleavage site. DNA recombination is performed outside the cell, but the resulting "recombinant" plasmid can be introduced into it by a process known as transformation and large quantities of the heterologous gene-containing recombinant plasmid are obtained by growing the transformant. Moreover, where the gene is properly inserted with reference to portions of the plasmid which govern the transcription and translation of the encoded DNA message, the resulting expression vehicle can be used to actually produce the polypeptide sequence for which the inserted gene codes, a process referred to as expression.

Expression is initiated in a region known as the promoter which is recognized by and bound by RNA polymerase. In some cases, as in the tryptophan or "trp" promoter preferred in the practice of the present invention, promoter regions are overlapped by "operator" regions to form a combined promoter-operator. Operators are DNA sequences which are recognized by so-called repressor proteins which serve to regulate the frequency of transcription initiation at a particular promoter. The polymerase travels along the DNA, transcribing the information contained in the coding strand from its 5' to 3' end into messenger RNA which is in turn translated into a polypeptide having the amino acid

sequence for which the DNA codes. Each amino acid is encoded by a nucleotide triplet or "codon" within what may for present purposes be referred to as the "structural gene", i.e. that part which encodes the amino acid sequence of the expressed product. After binding to the promoter, the RNA polymerase first transcribes nucleotides encoding a ribosome binding site, then a translation initiation or "start" signal (ordinarily ATG, which in the resulting messenger RNA becomes AUG), then the nucleotide codons within the structural gene itself. So-called stop codons are transcribed at the end of the structural gene whereafter the polymerase may form an additional sequence of messenger RNA which, because of the presence of the stop signal, will remain untranslated by the ribosomes. Ribosomes bind to the binding site provided on the messenger RNA, in bacteria ordinarily as the mRNA is being formed, and themselves produce the encoded polypeptide, beginning at the translation start signal and ending at the previously mentioned stop signal. The desired product is produced if the sequences encoding the ribosome binding site are positioned properly with respect to the AUG initiaton codon and if all remaining codons follow the initiaton codon in phase. The resulting product may be obtained by lysing the host cell and recovering the product by appropriate purification from other bacterial protein.

While isolation from donor fibroblasts has provided sufficient material for partial characterization and limited clinical studies with homogeneous leukocyte interferon, it is a totally inadequate source for the amounts of interferon needed for large scale clinical trials and for broad scale prophylactic and/or therapeutic use thereafter. Indeed, presently clinical investigations employing human fibroblast-derived interferons in antitumor and antiviral testing have principally been confined to crude (<1 percent pure) preparations of the material, and long lead times for the manufacture of sufficient quantities, even at unrealistic price levels, have critically delayed

investigation on an expanded front.

We perceived that application of recombinant DNA technology would be the most effective way of providing large quantities of fibroblast interferon which, despite the absence in material so produced of the glycosylation characteristic of human-derived material, could be employed clinically in the treatment of a wide range of viral and neoplastic diseases and have succeeded in producing mature human fibroblast interferon microbially, by constructing a gene therefor which could then be inserted in microbial expression vehicles and expressed under the control of microbial gene regulatory controls.

Our approach to obtaining a fibroblast gene involved the following tasks:

1. Partial amino acid sequences of human fibroblast interferon were used to construct sets of synthetic DNA probes the codons of which, in the aggregate, represented all the possible combinations capable of encoding the partial amino acid sequences.

2. Bacterial colony banks were prepared containing complementary DNA (cDNA) from induced messenger RNA. The probes of part (1) were used to prime the synthesis of radio-labelled single stranded cDNA for use as hybridization probes. The synthetic probes would hybridize with induced mRNA as template and be extended by reverse transcription to form induced, radio-labelled cDNA. Clones from the colony bank that hybridized to radio-labelled cDNA obtained in this manner have been investigated further to confirm the presence of a full-length interferon encoding gene. Any partial length putative gene fragment obtained was itself used as a probe for the full-length gene.

3. The full-length gene obtained above was tailored, using synthetic DNA, to eliminate any leader sequence that might prevent microbial expression of the mature polypeptide and to permit appropriate positioning in an expression vehicle relative to start signals and the ribosome binding site of a microbial promoter. Expressed interferon was purified to a point permitting confirmation of its character and determination of its activity.

In applying methods of recombinant DNA technology as outlined above a series of replicable plasmidic expression vehicles have been constructed which direct the high level synthesis in transformant microorganisms of a mature polypeptide with the properties of authentic human fibroblast interferon. The product polypeptide exhibits the amino acid sequence of such interferon and is active in in vitro testing despite the lack of glycosylation characteristic of the human-derived material. Reference herein to the "expression of mature fibroblast interferon" connotes the bacterial or other microbial production of an interferon molecule containing no glycosyl groups or a presequence that immediately attends mRNA translation of the human fibroblast interferon genome. Mature fibroblast interferon, according to the present invention, is immediately expressed from a translation start signal (ATG) which also encodes the first amino acid codon of the natural product. The presence or absence of the methionine first amino acid in the microbially expressed product is governed by a kinetic phenomenon dependent on fermentation growth conditions and/or levels of expression in the transformant host. Mature fibroblast interferon could be expressed together with a conjugated protein other than the conventional leader, the conjugate being specifically cleavable in an intra- or extracellular environment (see British Patent Publication No. 2007676A). Finally, the mature interferon could be produced in conjuction with a microbial "signal" peptide which transports the conjugate to the cell wall, where the signal is processed away and the mature poly-

peptide secreted.

Figures 1 to 5 appended hereto are described in the detailed text infra. Figure 6 schematically depicts the construction of plasmids coding for the direct expression of mature fibroblast interferon. Restriction sites and residues are as shown ("Pst I",etc.). "Ap^R" and "Tc^R" connote portions of the plasmid which express, respectively, ampicillin and tetracycline resistance. The legend "p o" is an abbreviation for "promoter operator".

DESCRIPTION OF THE PREFERRED EMBODIMENTS

A. Microorganisms employed

The work described involved use of the microorganism E. coli K-12 strain 294 (end A, thi^-, hsr^-, hsm_k^+), as described in British Patent Publication No. 2055382 A. This strain has been deposited with the American Type Culture Collection, ATCC Accession No. 31446 on Oct.28,1978. All recombinant DNA work was performed in complicance with applicable guidelines of the National Institutes of Health.

The invention although described in its most preferred embodiments with reference to E. coli K-12 strain 294, defined above, comprises also other known E. coli strains such as E. coli B, E. coli x 1776 and E. coli W 3110, or other microbial strains many of which are deposited and (potentially) available from recognized microorganism depository institutions, such as the American Type Culture Collection (ATCC). See also German Offenlegungsschrift 2644432. These other microorganisms include, for example, Bacilli such as Bacillus subtilis and other enterobacteriaceae among which can be mentioned as examples Salmonella typhimurium and Serratia marcescens, utilizing plasmids that can replicate and express heterologous gene sequences therein. Yeast, such as Saccharomyces cerevisiae, may also be employed to advantage as host organism in the

preparation of the interferon protein hereof by expression of genes coding therefor under the control of a yeast promoter.

B. General methods

Restriction enzymes were purchased from New England Biolabs and used as directed. Plasmid DNA was prepared by a standard cleared lysate procedure (D.B. Clewell, J. Bacteriol. 110, 667-676 [1972]) and purified by column chromatography on Biogel A-50M. DNA sequencing was performed using the method of Maxam and Gilbert (Methods Enzymol. 65, 499-560 [1980]). DNA restriction fragments were isolated from polyacrylamide gels by electroelution. DNA fragments were radiolabeled for use as hydridization probes by the random calf thymus DNA priming procedure of Taylor et al. (Biochim. Biophys. Acta 442, 324-330 [1976]). In situ colony hybridizations were performed by the Grunstein-Hogness procedure (Proc. Natl. Acad. Sci. USA 72, 3961-3965 [1975]).

C. Chemical synthesis of deoxyoligonucleotides

The deoxyoligonucleotides were synthesized by the modified phosphotriester method in solution (Crea et al., Proc. Natl. Acad. Sci. USA 75, 5765-5769 [1978]), using trideoxynucleotides as building blocks (Hirose et al., Tetrahedron Letters 28, 2449-2452 [1978]). The materials and general procedures were similar to those described by Crea et al., Nucleic Acids Res. 8, 2331-2348 [1980]. The six pools of primers (Figure 1) containing four dodecanucleotides each were obtained by separately coupling two hexamer pools (of two different 5'-terminal sequences each) with three different hexamer pools (of two different 3'-terminal sequences each).

D. Induction of fibroblasts

Human fibroblasts (cell line GM-2504A) were grown as described previously by Pestka et al., Proc. Natl. Acad. Sci. USA 72, 3898-3901 [1975]. Growth medium (Eagle's minimal essential medium containing 10% fetal calf serum) was removed from roller bottles (850 cm^3) and replaced with 50 ml growth medium containing 50 µg/ml of poly(I): poly(C) and 10 µg/ml cycloheximide. This induction medium was removed after 4 hours at 37°C and cell monolayers were washed with phosphate buffered saline (PBS; 0.14M NaCl, 3mM KCl, 1.5 mM KH_2PO_4, 8mM Na_2HPO_4). Each bottle was incubated at 37°C with 10 ml of a trypsin - EDTA solution (Gibco 610-5305) until cells were detached, and fetal calf serum was added to a concentration of 10%. Cells were spun for 15 minutes at 500 x g and pellets were resuspended in PBS, pooled, and resedimented. Cells were frozen in liquid nitrogen. Approximately 0.17 g of cells were obtained per roller bottle.

E. Preparation and assay of interferon mRNA

Poly(A)-containing mRNA was prepared from human fibroblasts by phenol extractions and oligo(dT)-cellulose chromatography as described by Green et al. (Arch. Biochem. Biophys. 172, 74-89 [1975]). The poly (A) containing RNA was enriched for interferon mRNA by centrifugation on a linear 5-20% (w/v) sucrose gradient. The RNA samples were heated to 80°C for 2 minutes, rapidly cooled, layered over the gradient, and centrifuged for 20 hours at 30,000 rpm at 4°C in a Beckman SW-40 rotor. Fractions were collected, ethanol precipitated, and dissolved in H_2O.

One microgram samples of mRNA were injected into Xenopus laevis oocytes as described by Cavalieri et al., Proc. Natl. Acad. Sci. USA 74, 3287-3291 [1977]. The injected oocytes were incubated 24 hours at 21°C, homogenized, and centrifuged for 5 minutes at 10,000 x g. The

interferon in the supernatant was determined by the cytopathic effect (CPE) inhibition assay (Stewart, The Interferon System, Springer-Verlag, New-York-Wien, 1979) using Sindbis virus and human diploid cells (WISH). Interferon titers of 1,000 to 6,000 units recovered (NIH reference standard) per microgram of RNA injected were routinely obtained for the 12S species of mRNA.

F. Synthesis and cloning of cDNA

Single stranded cDNA was prepared in 100 µl reactions containing 5 µg of 12S fraction mRNA, 20 mM Tris-HCl (pH 8.3), 20 mM KCl, 8mM $MgCl_2$, 30 mM β-mercaptoethanol, 100 µCi of (α^{32}P)dCTP and 1 mM dATP, dCTP, dGTP, dTTP. The primer was the synthetic HindIII decamer dCCAAGCTTGG (Scheller et al., Science 196, 177-180 [1977]), which had been extended at the 3'-terminus with about 20 to 30 deoxythymidine residues using terminal deoxynucleotidyl transferase (Chang et al., Nature 275, 617-624 [1978]). 100 units of reverse transcriptase were added and the reaction mixture was incubated at 42°C for 30 minutes. The second strand DNA synthesis was carried out as described previously (Goeddel et al., Nature 281, 544-548 [1979]). The double stranded cDNA was treated with 1200 units of S1 nuclease for 2 hours at 37°C in 25 mM sodium acetate (pH 4.5), 1mM $ZnCl_2$, 0.3M NaCl. After phenol extraction the mixture was separated electrophoretically on a 8% polyacrylamide gel. cDNA (~ 0.5 µg) ranging from 550 to 1500 base pairs in size was recovered by electroelution. A 20 ng aliquot was extended with deoxyC residues using terminal deoxynucleotidyl transferase (Chang et al., supra), and annealed with 100 ng of pBR322 which had been cleaved with PstI and tailed with deoxyG residues (Chang et al., supra). The annealed mixture was used to transform E. coli K-12 strain 294 by a published procedure (Hershfield et al., Proc. Natl. Acad. Sci. USA 71, 3455-3459 [1974]).

G. Preparation of induced and uninduced ^{32}P-cDNA probes

5 µg of 12S mRNA were combined with either 2 µg of oligo $(dT)_{12-18}$ or 5 µg of each synthetic primer pool (Figure 1) in 60 µl of 10mM Tris-HCl (pH 8), 1 mM EDTA. The mixtures were boiled 3 minutes, and quenched on ice. 60 µl of 40 mM Tris-HCl (pH 8.3), 40 mM KCl, 16mM $MgCl_2$, 60 mM β-mercaptoethanol, 1 mM dATP, dGTP, dTTP and 5 x 10^{-7}M (α-^{32}P) dCTP (2,000 - 3,000 Ci/mM) was added to each template-primer mixture at 0°C. After the addition of 100 units of reverse transcriptase, the reactions were incubated at 42°C for 30 minutes and purified by passage over 10 ml Sephadex G-50 columns. The products were treated with 0.3N NaOH for 30 minutes at 70°C, neutralized, and ethanol precipitated.

The ^{32}P-cDNAs were combined with 100 µg of poly(A) mRNA from uninduced fibroblasts in 50 µl of 0.4M sodium phosphate (pH 6.8), 0.1% sodium dodecyl sulfate (SDS). The mixtures were heated at 98°C for 5 minutes and allowed to anneal 15 hours at 45°C. The DNA-RNA hybrids (containing uninduced cDNA sequences) were separated from single-stranded DNA (induced cDNA sequences) by chromatography on hydroxyapatite as described by Galau et al. (Proc. Natl. Acad. Sci. USA 74, 1020-1023 [1977]). The DNA-RNA hybrids were treated with alkali to remove RNA.

H. Screening of recombinant plasmids with ^{32}P-cDNA probes

Approximately 1 µg samples of plasmid DNA were prepared from individual transformants by a published procedure (Birnboim et al., Nucleic Acids Res. 7, 1513-1523 [1979]). The DNA samples were linearized by digestion with EcoRI, denatured in alkali, and applied to each of three nitrocellulose filters by the dot hybridization procedure (Kafatos et al., Nucleic Acids Res. 7, 1541-1552 [1979]). The filters were hybridized with the ^{32}P-cDNA probes for 16 hours at 42°C in 50% formamide, 10x Denhardt's solution

(Biochem. Biophys. Res. Comm. 23, 641-646 [1966]), 6xSSC, 40 mM Tris-HCl (pH 7.5), 2mM EDTA, 40 µg/ml yeast RNA. Filters were washed with 0.1xSSC, 0.1% SDS twice for 30 minutes at 42°C, dried, and exposed to Kodak XR-2 x-ray film using Dupont Lightning-Plus intensifying screens at -80°C. [SSC contains 0.15 M NaCl and 0.015 M sodium citrate, pH 7.0].

I. Construction of plasmids for direct expression of FIF

The synthetic primers I (dATGAGCTACAAC) and II (dCATGAGCTACAAC) were phosphorylated using T4 polynucleotide kinase and (γ-^{32}P)ATP to a specific activity of 700 Ci/mM as described by Goeddel et al., Proc. Natl. Acad. Sci. USA 76, 106-110 [1979]. Primer repair reactions were performed as follows: 250 pM of the ^{32}P-ppimers were combined with 8 µg (10 pM) of a 1200 bp HhaI restriction fragment containing the FIF cDNA sequence. The mixture was ethanol precipitated, resuspended in 50 µl H_2O, boiled 3 minutes, quenched in a dry ice-ethanol bath, and combined with a 50 µl solution of 20mM Tris-HCl (pH 7.5), 14 mM $MgCl_2$, 120 mM NaCl, 0.5 mM dATP, dCTP, dGTP, dTTP at 0°C. 10 units of DNA polymerase I Klenow fragment were added and the mixture was incubated at 37°C for 4 1/2 hours. Following extraction with phenol/$CHCl_3$ and restriction with PstI, the desired product was purified on a 6% polyacrylamide gel. Subsequent ligations were done at room temperature (cohesive termini) or 4°C (blunt ends) using conditions reported previously (Goeddel et al., supra).

J. Assay for interferon expression in E. coli

Bacterial extracts were prepared for IF assay as follows: One ml cultures were grown overnight in LB (Luria-Bertani) medium containing 5 µg/ml tetracycline, then diluted into 25 ml of M9 medium supplemented with 0.2% glucose, 0.5% casamino acids and 5 µg/ml tetracycline. 10 ml samples were harvested by centrifugation when

absorbance at 550 nm (A_{550}) reached 1.0. The cell pellets were quickly frozen in a dry ice-ethanol bath and cleared lysates were prepared as described by Clewell (supra). Interferon activity in the supernatants was determined by comparison with NIH FIF standards using CPE inhibition assays. Two different assays were used: (a) WISH (human amnion) cells were seeded in microtiter dishes. Samples were added 16 to 20 hours later and diluted by serial 2-fold dilution. Sindbis virus was added after at least 3 hours of incubation. Plates were stained 20 to 24 hours later with crystal violet. (b) MDBK (bovine kidney) cell line was seeded simultaneously with 2-fold dilutions of samples. Vesicular stomatitis virus was added after 2 to 3 hours incubation and plates were stained with crystal violet 16 to 18 hours later. To test pH 2 stability bacterial extracts and standards were diluted in minimal essential medium to a concentration of 1000 units/ml. One ml aliquots were adjusted to pH 2 with 1N HCl, incubated at 4°C for 16 hours, and neutralized by addition of NaOH. IF activity was determined by the CPE inhibition assay using human amnion cells. To establish antigenic identity 25 µl aliquots of the 1000 U/ml interferon samples (untreated) were incubated with 25 µl of rabbit antihuman leukocyte interferon for 60 minutes at 37°C, centrifuged at 12,000 x g for 5 minutes and the supernatant assayed. Fibroblast and leukocyte interferon standards were obtained from the National Institutes of Health. Rabbit antihuman leukocyte interferon was obtained from the National Institute of Allergy and Infectious Diseases.

K. Chemical synthesis of primer pools complementary to FIF mRNA

The known amino-terminal protein sequence of human fibroblast interferon permitted us to deduce the 24 possible mRNA sequences which could code for the first four amino acids. The 24 complementary deoxyoligonucleotides were synthesized in 6 pools of 4 dodecamers each (Figure 1).

The six pools of 4 deoxyoligonucleotides each were synthesized by the modified phosphotriester method in solution and on solid phase (Crea et al., supra). The basic strategy involved reacting two different 3'-blocked trimers with an excess of a single 5'-protected trimer to yield a pool of two hexamers, each represented equally. The coupling of two pools, each containing two hexamers, then resulted in a pool of four dodecamers.

L. Identification of FIF cDNA clones

Using 12S mRNA from induced human fibroblasts (1,000 units IF activity per µg in oocyte assay), double stranded cDNA was prepared and inserted into pBR322 at the PstI site by the standard dG:dC tailing method as described by Chang et al., supra. A fibroblast cDNA library consisting of 30,000 ampicillin-sensitive, tetracycline-resistant transformants of E. coli K-12 strain 294 was obtained from 20 ng of cDNA ranging in size from 550 to 1300 base pairs. Plasmid DNA was prepared from 600 of the transformants and applied to 3 sets of nitrocellulose filters as described above.

The approach followed in the identification of hybrid plasmids containing fibroblast interferon cDNA sequences was similar to that used to identify human leukocyte interferon recombinant plasmids (Goeddel et al., Nature 287, 411-416 [1980]). Radiolabeled cDNA hybridization probes were prepared using either the 24 synthetic dodecamers or oligo$(dT)_{12-18}$ as primers and 12S RNA from induced fibroblasts (5000 units/µg in oocytes) as template. The ^{32}P-cDNAs (specific activity $> 5 \times 10^8$ cpm/µg) obtained were hybridized to a large excess of mRNA isolated from uninduced human fibroblasts, and the mRNA-cDNA hybrids were separated from unreacted cDNA by hydroxyapatite chromatography (Galau et al., supra). The single stranded cDNA fractions should be enriched for sequences which are present in induced fibroblasts but absent in uninduced

cells, and the mRNA-cDNA hybrids should represent sequences common to both induced and uninduced cells. Approximately 4×10^6 cpm of single stranded cDNA (hybridization probe A) and 8×10^6 cpm of cDNA-mRNA hybrids were obtained using oligo$(dT)_{12-18}$ primed cDNA; 1.5×10^6 cpm of single stranded (hybridization probe B) and 1.5×10^6 cpm of hybrids were obtained from cDNA primed using synthetic dodecamer pools 1-6. The cDNA-mRNA hybrids from both fractionations were combined, the RNA hydrolyzed by treatment with alkali, and the ^{32}P-cDNA used as hybridization probe C. Many of the 600 plasmid samples hybridized with both probes A and C, indicating that the hybridization reactions between uninduced mRNA and ^{32}P-cDNA (prior to the hydroxyapatite fractionation step) had not gone to completion. However, only one of the 600 plasmids (pF526) hybridized strongly with the specifically primed, induced cDNA probe B (Figure 2). Plasmid pF526 also hybridized with the total oligo$(dT)_{12-18}$ primed, induced cDNA probe A, and failed to give detectable hybridization to the combined uninduced probe C.

PstI digestion of pF526 showed the cloned cDNA insert to be about 550 base pairs long, probably too short to contain the entire coding region for fibroblast interferon. Therefore, a ^{32}P-labeled DNA probe was prepared from this PstI fragment by random priming with calf thymus DNA (Taylor et al., supra). This probe was used to screen 2000 individual colonies from a newly constructed fibroblast cDNA library (the new cDNA library was prepared using 12S mRNA from induced fibroblasts having a titer of 6,000 units/ml in the oocyte assay system). Sixteen clones hybridized to the probe. Plasmids prepared from the majority of these released two fragments when cleaved with PstI, indicating that the cDNA contained an internal PstI site. Clone pFIF3 contained the largest cDNA inert, about 800 base pairs. The DNA sequence of the insert was determined by the Maxam-Gilbert procedure (supra) and is shown in Figure 3. The amino acid sequence of human fibro-

blast interferon predicted from the nucleotide sequence is identical to that reported recently by Taniguchi et al. (Gene 10, 11-15 [1980]) and by Derynck et al. (supra) from DNA sequencing of FIF cDNA clones. A precursor or signal peptide of 21 amino acids is followed by a sequence of 166 amino acids representing the mature interferon, a stretch of 196 3'-untranslated nucleotides and a poly(A) tail. The NH_2-terminal 20 amino acids of mature FIF have been direclty determined by protein microsequencing and are the same as those predicted from the DNA sequence.

M. Direct expression of fibroblast interferon

To express high levels of mature fibroblast interferon in E. coli initiation of protein synthesis must occur at the ATG codon of the mature polypeptide (amino acid 1) rather than at the ATG of the signal peptide (amino acid S1) (Figure 3).

Our approach to removing the signal peptide coding regions from pFIF3 is depicted in Figure 4. A 1200 bp DNA fragment which contained the entire cDNA insert was isolated from a polyacrylamide gel after digesting pFIF3 with HhaI. Two separate synthetic deoxyoligonucleotide primers, dATGAGCTACAAC(I) and dCATGAGCTACAAC(II), were prepared. Both primers contain the coding sequence for the first four amino acids of mature fibroblast interferon; primer II has an additional C at the 5'-terminus. Primer repair reactions and subsequent ligations were carried out separately for primers I and II, and gave nearly identical results. Therefore, only reactions using primer I are discussed in detail here. The primers were 5'-radiolabeled using (γ-^{32}p)ATP and T4 polynucleotide kinase, combined with the 1200 bp HhaI DNA fragment and the mixtures were denatured by boiling. Following hybridization of the primer to the denatured HhaI DNA fragment, E. coli DNA polymerase I Klenow fragment (Klenow et al.,

Proc. Natl. Acad. Sci. USA 65, 168-175 [1970]) was used to catalyze the repair synthesis of the plus (top) strand (Figure 4). In addition, the associated 3' → 5' exonuclease activity of the Klenow fragment removed the 3'-protruding end from the minus (bottom) strand, leaving a flush end. Analysis of samples of the reaction mixture by polyacrylamide gel electrophoresis indicated that the repair synthesis did not go to completion, but stopped at several discrete sites. Therefore, the entire reaction mixture was treated with PstI and the desired 141 bp fragment (180,000 Cerenkow cpm; ~0.3 pM) was purified by polyacrylamide gel electrophoresis (Figure 5). Ligation of this fragment to 1 µg (~4 pM) of the 363 bp PstI-BglII fragment isolated from pFIF3 (Fig. 4), followed by BglII digestion, yielded 50,000 Cerenkov cpm (~0.1 pM, ~30 ng) of the 504 bp DNA fragment containing the entire coding sequence for mature fibroblast interferon. The same reactions using primer II gave 83,000 cpm (~0.15 pM, ~50 ng) of 505 bp product.

The construction of plasmids which direct the synthesis of human fibroblast interferon is outlined in Figure 6. Separate expression plasmids were constructed which placed FIF synthesis under the control of the E. coli lac or trp promoter-operator systems. Both ot these systems have proven useful for the direct expression of eukaryotic genes in E. coli: human growth hormone has been efficiently synthesized using the lac system (Goeddel et al., Nature 281, 544-548 [1979]) and human leukocyte interferon has been produced at high levels using the trp system (Goeddel et al., Nature 287, 411 [1980]).

pBRH trp was digested with EcoRI restriction enzyme and the resulting fragment isolated by PAGE and electroelution. EcoRI-digested plasmid pSom 11 (Itakura et al., Science 198, 1056-1063 [1977]); G.B. patent publication no. 2 007 676 A) was combined with the above fragment. The mixture was ligated with T_4 DNA ligase and the resul-

ting DNA transformed into E. coli K-12 strain 294 as previously described. Transformant bacteria were selected on amicillin-containing plates. Resulting ampicillin-resistant colonies were screened by colony hybridization (Grunstein et al., supra) using as a probe the trp promoter-operator containing the above fragment isolated from pBRHtrp, which had been radioactively labelled with P^{32}. Several colonies shown positive by colony hybridization were selected, plasmid DNA was isolated and the orientation of the inserted fragments determined by restriction analysis employing restriction enzymes BglII and BamHI in double digestion. E. coli 294 containing the plasmid designated pSOM7Δ2, which has the trp promoter-operator fragment in the desired orientation was grown in LB medium containing 10 µg/ml ampicillin. The cells were grown to optical density 1 (at 550 nM), collected by centrifugation and resuspended in M9 media in tenfold dilution. Cells were grown for 2-3 hours, again to optical density 1, then lysed and total cellular protein analyzed by SDS urea (15%) PAGE (Maizel et al., Methods Virol. 5, 180-246 [1971]).

Plasmid pBR322 was HindIII digested and the protruding HindIII ends in turn digested with S1 nuclease. The S1 nuclease digestion involved treatment of 10 µg of HindIII-cleaved pBR322 in 30 µl S1 buffer (0.3 M NaCl, 1 mM $ZnCl_2$, 25 mM sodium acetate, pH 4.5) with 300 units S1 nuclease for 30 minutes at 15°C. The reaction was stopped by the addition of 1 µl of 30 x S1 nuclease stop solution (0.8M Tris base, 50 mM EDTA). The mixture was phenol extracted, chloroform extracted and ethanol precipitated, then EcoRI digested as previously described and the large fragment (1) obtained by PAGE procedure followed by electroelution. The fragment obtained has a first EcoRI sticky end and a second, blunt end whose coding strand begins with the nucleotide thymidine.

Plasmid pSom7Δ2, as prepared above, was BglII digested and the BglII sticky ends resulting made double stranded with the Klenow polymerase I procedure using all four deoxynucleotide triphosphates. EcoRI cleavage of the resulting product followed by PAGE and electroelution of the small fragment (2) yielded a linear piece of DNA containing the tryptophan promoter-operator and codons of the LE' "proximal" sequence upstream from the BglII site ("LE'(p)"). The product had an EcoRI end and a blunt end resulting from filling in the BglII site. However, the BglII site is reconstituted by ligation of the blunt end of fragment (2) to the blunt end of fragment (1). Thus, the two fragments were ligated in the presence of T_4 DNA ligase to form the recirculated plasmid pHKY 10 which was propagated by transformation into competent E. coli strain 294 cells.

Plasmid pGM1 carries the E. coli tryptophan operon containing the deletion ΔLE1413 (Miozzari et al., J. Bacteriology 133, 1457-1466 [1978]) and hence expresses a fusion protein comprising the first 6 amino acids of the trp leader and approximately the last third of the trp E polypeptide (hereinafter referred to in conjunction as LE'), as well as the trp D polypeptide in its entirety, all under the control of the trp promoter-operator system. The plasmid, 20 µg, was digested with the restriction enzyme PvuII which cleaves the plasmid at five sites. The gene fragments were next combined with EcoRI linkers (consisting of a self complementary oligonucleotide of the sequence: pCATGAATTCATG) providing an EcoRI cleavage site for a later cloning into a plasmid containing an EcoRI site. The 20 µg of DNA fragments obtained from pGM1 were treated with 10 units T_4 DNA ligase in the presence of 200 pico moles of the 5'-phosphorylated synthetic oligonucleotide pCATGAATTCATG and in 20 µl T_4 DNA ligase buffer (20mM Tris, pH 7.6, 0.5 mM ATP, 10 mM $MgCl_2$, 5 mM dithiothreitol) at 4°C overnight. The solution was then heated 10 minutes at 70°C to halt ligation. The linkers

were cleaved by EcoRI digestion and the fragments, now with EcoRI ends were separated using 5% PAGE and the three largest fragments isolated from the gel by first staining with ethidium bromide, locating the fragments with ultraviolet light, and cutting from the gel the portions of interest. Each gel fragment, with 300 microliters 0.1xTBE, was placed in a dialysis bag and subjected to electrophoresis at 100 V for one hour in 0.1xTBE buffer (TBE buffer contains: 10.8 gm Tris base, 5.5 gm boric acid, 0.09 gm Na_2EDTA in 1 liter H_2O). The aqueous solution was collected from the dialysis bag, phenol extracted, chloroform extracted and made 0.2 M sodium chloride, and the DNA recovered in water after ethanol precipitation. The trp promoter-operator containing gene with EcoRI sticky ends was identified in the procedure next described, which entails the insertion of fragments into a tetracycline sensitive plasmid which, upon promoter-operator insertion, becomes tetracycline resistant.

Plasmid pBRH1 (Rodriguez et al., Nucleic Acids Research 6, 3267-3287 [1979]) expressed ampicillin resistance and contains the gene for tetracycline resistance but, there being no associated promoter, does not express that resistance. The plasmid is accordingly tetracycline sensitive. By introducing a promoter-operator system in the EcoRI site, the plasmid can be made tetracycline resistant.

pBRH1 was digested with EcoRI and the enzyme removed by phenol extraction followed by chloroform extraction and recovered in water after ethanol precipitation. The resulting DNA molecule was, in separate reaction mixtures, combined with each of the three DNA fragments obtained above and ligated with T_4 DNA ligase as previously described. The DNA present in the reaction mixture was used to transform competent E. coli K-12 strain 294 by standard techniques (Hershfield et al., supra) and the bacteria

plated on LB plates containing 20 µg/ml ampicillin and 5 µg/ml tetracycline. Several tetracycline-resistant colonies were selected, plasmid DNA isolated and the presence of the desired fragment confirmed by restriction enzyme analysis. The resulting plasmid is designated pBRHtrp.

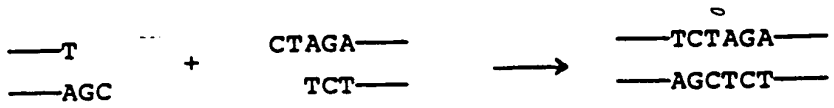
An EcoRI and BamHI digestion product of the viral genome of hepatitis B was obtained by conventional means and cloned into the EcoRI and BamHI sites of plasmid pGH6 (Goeddel et al., Nature 281, 544-548 [1979]) to form the plasmid pHS32. Plasmid pHS32 was cleaved with XbaI, phenol extracted, chloroform extracted and ethanol precipitated. It was then treated with 1 µl E. coli polymerase I, Klenow fragment, in 30 µl polymerase buffer (50 mM potassium phosphate pH 7.4, 7mM $MgCl_2$, 1 mM β-mercaptoethanol) containing 0.1mM dTTP and 0.1mM dCTP for 30 minutes at 0°C then 2 hours at 37°C. This treatment causes 2 of the 4 nucleotides complementary to the 5' protruding end of the XbaI cleavage site to be filled in:

```
5'  CTAGA——                    5'  CTAGA——
3'      T——        ⟶           3'    TCT——
```

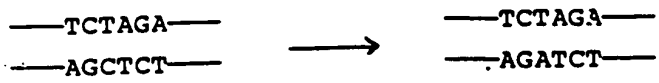
Two nucleotides, dC and dT, were incorporated giving an end with two 5' protruding nucleotides. This linear residue of plasmid pHS32 (after phenol and chloroform extraction and recovery in water after ethanol precipitation) was cleaved with EcoRI. The large plasmid fragment was separated from the smaller EcoRI-XbaI fragment by PAGE and isolated after electroelution. This DNA fragment from pHS32 (0.2 µg) was ligated under conditions similar to those described above, to the EcoRI-Taq I fragment of the tryptophan operon (~0.01 µg), derived from pBRHtrp.

In the process of ligating the fragment from pHS32 to the EcoRI-TaqI fragment, as described above, the

TaqI protruding end is ligated to the XbaI remaining protruding end even though it is not completely Watson-Crick base-paired:



A portion of this ligation reaction mixture was transformed into E. coli 294 cells, heat treated and plated on LB plates containing ampicillin. Twenty-four colonies were selected, grown in 3 ml LB media, and plasmid isolated. Six of these were found to have the XbaI site regenerated via E. coli catalyzed DNA repair and replication.



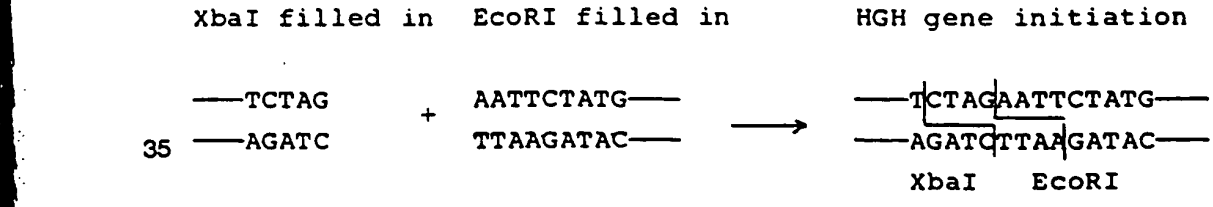
These plasmids were also found to cleave both with EcoRI and HpaI and to give the expected restriction fragments. One plasmid, designated pTrp14, was used for expression of heterologous polypeptides, as next discussed.

The plasmid pHGH 107 (Goeddel et al., Nature 281, 544-548 [1979]) contains a gene for human growth hormone made up of 23 amino acid codons produced from synthetic DNA fragments and 163 amino acid codons obtained from complementary DNA produced via reverse transcription of human growth hormone messenger RNA. This gene, though it lacks the codons of the "pre" sequence of human growth hormone, does contain an ATG translation initiation codon. The gene was isolated from 10 µg pHGH 107 after treatment with EcoRI followed by E. coli polymerase I Klenow fragment and dTTP and dATP as described above. Following phenol and chloroform extraction and ethanol precipitation the plasmid was treated with BamHI.

The human growth hormone (HGH) gene-containing fragment was isolated by PAGE followed by electroelution. The resulting DNA fragment also contains the first 350 nucleotides of the tetracycline resistance structural gene, but lacks the tetracycline promoter-operator system so that, when subsequently cloned into an expression plasmid, plasmids containing the insert can be located by the restoration of tetracycline resistance. Because the EcoRI end of the fragment has been filled in by the Klenow polymerase I procedure, the fragment has one blunt and one sticky end, ensuring proper orientation when later inserted into an expression plasmid.

The expression plasmid pTrp14 was next prepared to receive the HGH gene-containing fragment prepared above. Thus, pTrp14 was XbaI digested and the resulting sticky ends filled in with the Klenow polymerase I procedure employing dATP,dTTP, dGTP and dCTP. After phenol and chloroform extraction and ethanol precipitation the resulting DNA was treated with BamHI and the resulting large plasmid fragment isolated by PAGE and electroelution. The pTrp14-derived fragment had one blunt and one sticky end, permitting recombination in proper orientation with the HGH gene containing fragment previously described.

The HGH gene fragment and the pTRP14 ΔXba-BamHI fragment were combined and ligated together under conditions similar to those described above. The filled in XbaI and EcoRI ends ligated together by blunt end ligation to recreate both the XbaI and the EcoRI site:



This construction also recreates the tetracycline resistance gene. Since the plasmid pHGH 107 expresses tetracycline resistance from a promoter lying upstream from the HGH gene (the lac promoter), this construction, designated pHGH 207, permits expression of the gene for tetracycline resistance under the control of the tryptophan promoter-operator. Thus the ligation mixture was transformed into E. coli 294 and colonies selected on LB plates containing 5 µg/ml tetracycline.

Plasmid pHGH 207 was EcoRI digested and the trp promoter containing EcoRI fragment recovered by PAGE followed by electroelution. Plasmid pBRH1 was EcoRI digested and the cleaved ends treated with bacterial alkaline phosphatase (BAP, 1 µg, in 50 mM Tris, pH 8, and 10 mM $MgCl_2$ for 30 min. at 65°C) to remove the phosphate groups on the protruding EcoRI ends. Excess bacterial alkaline phosphatase was removed by phenol extraction, chloroform extraction and ethanol precipitation. The resulting linear DNA, because it lacks phosphates on the protruding ends thereof, will in ligation accept only inserts whose complementary sticky ends are phosphorylated but will not itself recircularize, permitting more facile screening for plasmids containing the inserts.

The EcoRI fragment derived from pHGH 207 and the linear DNA obtained from pBRH1 were combined in the presence of T_4 ligase as previously described and ligated. A portion of the resulting mixture was transformed into E. coli strain 294 as previously described, plated on LB media containing 5 µg/ml of tetracycline, and 12 tetracycline resistant colonies selected. Plasmid was isolated from each colony and examined for the presence of a DNA insert by restriction endonuclease analysis employing EcoRI and XbaI: One plasmid containing the insert was designated pHKY1.

The plasmid pHKY10, described above, is a derivative of pBR322 which contains a BglII site between the tetracycline resistance (Tc^R) promoter and structural gene. The large DNA fragment isolated after digesting pHKY10 with PstI and BglII therefore contains part of the ampicillin resistance (Ap^R) gene and all of the Tc^R structural gene, but lacks the Tc^R promoter (Fig. 6). The plasmid pGH6 (Goeddel et al., Nature 281, 544-548 [1979]) was digested with EcoRI, the resulting single stranded ends were filled in with DNA polymerase I, and the plasmid was cleaved with PstI. The small fragment, containing part of the Ap^R gene, a double lac promoter and lac ribosome binding site, but lacking an ATG initiation triplet was isolated. A similar trp promoter fragment, containing the trp leader ribosome binding site, but lacking an ATG sequence (Goeddel et al., Nature 287, 411-416 [1980]), may be isolated from pHKY1 described above.

The trp fragment just referred to is an analog of the E. coli tryptophan operon from which the so-called trp attenuator has been deleted (Miozzari et al., J. Bact. 133, 1457-1466 [1978]) to controllably heighten expression levels. Expression plasmids containing the modified trp regulon can be grown to predetermined levels in nutrient media containing additive tryptophan in quantities sufficient to repress the promoter-operator system, then be deprived of tryptophan so as to derepress the system and occasion the expression of the intended product.

The expression plasmids may be assembled via three part ligation reactions as shown in Figure 6. 15 ng (0.05 pM) of the assembled FIF gene (504 or 505 bp), 0.5 µg (0.2 pM) of the large PstI - BglII fragment of pHKY10 and 0.2 µg (0.3 pM) of the appropriate promoter fragment were ligated and the mixture used to transform E. coli 294 (Goeddel et al., Nature 287, 411-416 [1980]). Plasmid DNA was prepared from individual transformants and analyzed by restriction mapping. Correct joining of the assembled gene to the promoter fragment should restore

the EcoRI (lac) or the XbaI (trp) recognition sequences. The majority of the plasmids gave the expected restriction enzyme digestion patterns. Individual clones (12 containing the trp promoter and 12 containing the lac promoter) were grown and extracts prepared for interferon assay as described above.

When assayed on human amnion (WISH) cells for antiviral activity by the CPE inhibition assay five of the trp transformants were positive (each approximately equivalent); eleven of the lac transformants gave equivalent IF activities. Therefore, one transformant from each series (pFIFlac9 and pFIFtrp69) was selected for further study (Table 1). DNA sequence analysis demonstrated that the desired attachment of promoter to FIF structural gene had occurred in both cases.

Table 1. Interferon activity in extracts of E. coli

E. coli K-12 strain 294 transformed by	Cell density (cells/ml)	IF Activity (units/l culture)	FIF molecules per cell
pBR322	3.5×10^8	-	-
pFIFlac9	3.5×10^8	9.0×10^6	2,250
pFIFtrp69	3.5×10^8	1.8×10^7	4,500
pFIFtrp^{3}69	3.5×10^8	8.1×10^7	20,200

Cells were grown and extracts prepared as described above. The human amnion (WISH) cell line was used for the CPE inhibition assay. Activities given are the average from three independent experiments. To determine the number of IF molecules per cell a FIF specific activity of 4×10^8 units/mg was used (Knight, supra).

The amounts of fibroblast interferon produced by pFIFlac9 and pFIFtrp69 are shown in Table 1. The trp

promoter gave a FIF expression level measurable higher than did the lac promoter. In an attempt to further increase FIF expression levels, pFIFtrp69 was cleaved with EcoRI and two 300 base pair EcoRI fragments containing the trp promoter (Goeddel et al., Nature 287, 411-416 [1980]) were inserted. The resulting plasmid, pFIFtrp^{3}69, contains three successive trp promoters which read toward the FIF gene. The amount of FIF synthesized by E. coli K-12 strain 294/pFIF trp^{3}69 is 4-5 times that produced by pFIFtrp69 (Table 1). This is apparently due to the derepression of the trp promoter which occurs when trp repressor levels are titrated by the multiple copies of the trp operator.

The FIF produced by E. coli K-12 strain 294/pFIFtrp69 behaves like authentic human FIF. As shown in Table 2, its antiviral activity is about 30 times greater on human cells than on bovine cells. In addition, the bacterially produced FIF is stable to treatment at pH 2 overnight and is not neutralized by rabbit antihuman leukocyte interferon antibodies (Table 3).

Table 2. Interferon activities measured on different cell types

Cells	Interferon Activity (units/ml)		
	LeIF	FIF	E. coli K-12 strain 294/pFIFtrp69 extract
Human amnion	20,000	10,000	1280
Bovine kidney	13,000	400	40

LeIF and FIF were NIH standard solutions having 20,000 units/ml and 10,000 units/ml respectively. Assays were performed as described above.

Table 3. Comparison of activities of extracts from E. coli K-12 strain 294/pFIFtrp69 with standard human leukocyte and fibroblast interferons

	Interferon Activity (units/ml)		
	LeIF	FIF	E. coli K-12 strain 294/pFIFtrp69
untreated	1000	1000	1000
pH 2	1000	1000	1000
rabbit antihuman LeIF antibodies	<16	1000	1000

Experimental procedures described above. Assayed by CPE inhibition using WISH cells/Sindbis virus.

N. Purification

The purification procedure for bacterial derived fibroblast interferon is as follows:

1. Frozen cells are suspended in twelve times volume per weight with sucrose lysis buffer (100mM Tris-HCl, 10% sucrose, 0.2M NaCl, 50mM EDTA, 0.2mM PMSF [phenylmethylsulfonyl chloride], pH 7.9) containing lysozyme at 1mg/ml. The cell suspension is stirred for 1 hour at 4°C and centrifuged. Fibroblast interferon activity remains in the supernatant.

2. Polyethyleneimine (5%, v/v) is added to the sonicated supernatant to a final concentration of 0.5% (v/v). The solution is stirred for 1 hour at 4°C and centrifuged. Interferon activity remains in the supernatant.

3. Solid ammonium sulfate is added to the polyethyleneimine supernatant to a final concentration of 50% saturation, stirred for 30 minutes at 4°C and centrifuged. Interferon activity is in the 50% pellet.

4. The 50% ammonium sulfate pellet is suspended in one half the volume of the 50% ammonium sulfate suspension with PBS (20 mM sodium phosphate, 0.15M NaCl, pH 7.4). Polyethylene glycol 6000 (50%, w/v, in PBS) is added to a final concentration of 12.5% (v/v), stirred at 4°C for 2 hours and centrifuged. Interferon activity is in the pellet. The pellet is suspended in a minimal volume of sucrose lysis buffer and clarified by centrifugation.

This initial extraction procedure results in a purification of fibroblast interferon from 0.001% of the total protein to 0.05% of the total protein. This material can be further purified to homogeneity by the following column chromatography steps:

5. Afinity chromatography on Amicon Blue B in sucrose lysis buffer.

6. Anion exchange chromatography on QAE Sephadex in sucrose lysis buffer in the absence of 0.2M NaCl.

7. Size exclusion chromatography on Sephadex G-75 in sucrose lysis buffer.

8. Reverse phase high pressure liquid chromatography.

O. Parenteral Administration

FIF may be parenterally administered to subjects requiring antitumor or antiviral treatment. Dosage and dose rate may parallel that currently in use in clinical investigations of human derived materials, e.g., about $(1\text{-}10) \times 10^6$ units daily, and in the case of materials of

purity greater than 1%, likely up to, e.g., 15 x 10^7 units daily. Dosages of bacterially obtained FIF could be significantly elevated for greater effect owing to the essential absence of human proteins other than FIF, which proteins in fibroblast-derived materials may act as pyrogens, exhibiting adverse effects, e.g., malaise, temperature elevation, etc.

As one example of an appropriate dosage form for essentially homogeneous bacterial FIF in parenteral form, 3 mg FIF of specific activity of, say, 2 x 10^8 U/mg may be dissolved in 25 ml of 5% human serum albumin, the solution is passed through a bacteriological filter and the filtered solution aseptically subdividied into 100 vials, each containing 6 x 10^6 units pure interferon suitable for parenteral administration. The vials are preferably stored in the cold (-20°C) prior to use.

The compounds of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby the polypeptide hereof is combined in admixture with a pharmaceutically acceptable carrier vehicle. Suitable vehicles and their formulation are described in Remington's Pharmaceutical Sciences by E.W. Martin, which is hereby incorporated by reference. Such compositions will contain an effective amount of the interferon protein hereof together with a suitable amount of vehicle in order to prepare pharmaceutically acceptable compositions suitable for effective administration to the host. One preferred mode of administration is parenteral.

What we claim is:

1. A polypeptide comprising the amino acid sequence of a mature human fibroblast interferon, microbially produced and unaccompanied by any corresponding presequence or portion thereof.

2. A polypeptide according to claim 1, unaccompanied by associated glycosylation.

3. The polypeptide according to claim 1, optionally containing the amino acid methionine as the ordinarily first amino acid of said interferon.

4. The polypeptide according to claim 1, optionally containing a cleavable conjugate or microbial signal protein attached to the N-terminus of the ordinarily first amino acid of said interferon.

5. A DNA sequence comprising a sequence coding for the polypeptide according to claims 1, 3 or 4.

6. The DNA sequence according to claim 5 operably linked with a DNA sequence capable of effecting microbial expression of a polypeptide according to claims 1, 3 or 4.

7. A replicable microbial expression vehicle capable, in a transformant microorganism, of expressing a polypeptide according to claims 1, 3 or 4.

8. A microbial expression vehicle according to claim 7 which is a plasmid.

9. A plasmid selected from the group consisting of pFIFlac9, pFIFtrp69, and pFIFtrp^{3}69.

10. A microorganism transformed with an expression vehicle according to any one of claims 7-9.

11. The microorganism according to claim 10, obtained by transforming an E. coli strain.

12. A transformed microorganism according to claim 11 wherein said E. coli strain is E. coli K-12 strain 294.

13. A transformed microorganism according to claim 10 obtained by transforming Bacillus subtilis.

14. A transformed microorganism according to claim 10 obtained by transforming Saccharomyces cerevisiae.

15. A composition of matter comprising a therapeutically active fraction of a polypeptide consisting essentially of the amino acid sequence of a mature human fibroblast interferon, the balance of said composition comprising soluble microbial protein from which said polypeptide may be purified to a degree sufficient for effective therapeutic application.

16. A bacterial extract comprising greater than about 95% pure polypeptide consisting essentially of the amino acid sequence of a mature fibroblast interferon according to any one of claims 1-4.

17. A pharmaceutical composition comprising a therapeutically effective amount of a mature human fibroblast interferon according to cla[illegible] 3 or 4, suitable for pharmaceutical administration.

18. The composition according to claim 17 suitable for parenteral administration.

19. A culture of microbial cells capable of producing a human fibroblast interferon in mature form.

20. The use of a mature human fibroblast interferon according to claims 1, 3 or 4, for antitumor or antiviral treatment or for preparing pharmaceutical compositions useful for such treatment.

21. A process for producing a polypeptide claimed in any one of claims 1-4 which process comprises causing a microorganism, transformed with a replicable microbial expression vehicle capable of expressing said polypeptide, to grow up and express said polypeptide and recovering it.

22. A process for producing microorganisms capable of expressing a polypeptide claimed in any one of claims 1-4 which process comprises transforming a microorganism with a replicable microbial expression vehicle capable of expressing said polypeptide and cultivating the transformed microorganism.

23. A process for producing a replicable microbial expression vehicle capable in a transformant microorganism of expressing a polypeptide as claimed in any one of claims 1-4, which process comprises constructing a first DNA sequence coding for said polypeptide and operably linking said first DNA sequence with a second DNA sequence capable of effecting microbial expression of said first DNA sequence.

24. The process of claim 23 wherein said second DNA sequence comprises a multiple trp promoter-operator.

25. The products and processes for their preparation as hereinbefore described.

***

What we claim is: AUSTRIA

1. A process for producing a polypeptide comprising the amino acid sequence of a mature human fibroblast interferon unaccompanied by any corresponding presequence or portion thereof, which process comprises causing a microorganism transformed with a replicable microbial expression vehicle capable of expressing said polypeptide, to grow up and to express said polypeptide and recovering it.

2. A process as claimed in claim 1, wherein the polypeptide is unaccompanied by associated glycosylation.

3. A process as claimed in claim 1, wherein the polypeptide optionally contains the amino acid methionine as the ordinarily first amino acid of said fibroblast interferon.

4. A process as claimed in claim 1, wherein the polypeptide optionally contains a cleavable conjugate or microbial signal protein attached to the N-terminus of the ordinarily first amino acid of said fibroblast interferon.

5. A process for the production of microorganisms capable of producing a polypeptide comprising the amino acid sequence of a mature human fibroblast interferon unaccompanied by any corresponding presequence or portion thereof, which process comprises transforming a microorganism with a replicable microbial expression vehicle capable of expressing said polypeptide and cultivating the transformed microorganism.

6. A process for the production of a replicable microbial expression vehicle capable in a transformant microorganism of expressing a polypeptide comprising the amino acid sequence of a mature human fibroblast interferon unaccompanied by any corresponding presequence or portion

thereof which process comprises constructing a first DNA sequence coding for said polypeptide and operably linking said first DNA sequence with a second DNA sequence capable of effecting microbial expression of said first DNA sequence.

7. The process of claim 6 wherein said second DNA sequence comprises a multiple trp promoter-operator.

8. A process for the preparation of pharmaceutical compositions containing a microbially produced polypeptide comprising the amino acid sequence of a mature human fibroblast interferon unaccompanied by any corresponding presequence or portion thereof which process comprises mixing said polypeptide with non-toxic, inert, therapeutically compatible carriers and bringing the resulting mixture into a suitable pharmaceutical dosage form.

9. A pharmaceutical composition containing a microbially produced polypeptide comprising the amino acid sequence of a mature human fibroblast interferon unaccompanied by any corresponding presequence or portion thereof and a non-toxic, inert, therapeutically compatible carrier material.

10. The processes for the preparation of polypeptides, replicable microbial expression vehicles and microorganisms as hereinbefore described.

11. A DNA sequence comprising a sequence coding for the amino acid sequence of a mature human fibroblast interferon unaccompanied by any corresponding presequence or portion thereof.

12. A DNA sequence comprising a sequence coding for the amino acid sequence of a mature human fibroblast interferon the amino acid methionine optionally being the ordinarily first amino acid of said interferon.

13. A DNA sequence comprising a sequence coding for the amino acid sequence of a mature human fibroblast interferon as well as for the amino acid sequence of a cleavable conjugate or microbial signal protein attached to the N-terminus of the ordinarily first amino acid of said interferon.

14. A DNA sequence according to any one of claims 11-13 operably linked with a DNA sequence capable of effecting microbial expression of the encoded polypeptide.

15. A replicable microbial expression vehicle capable, in a transformant microorganism, of expressing a polypeptide coded by a DNA sequence according to any one of claims 11-14.

16. An expression vehicle as claimed in claim 15 which is a plasmid.

17. A plasmid selected from the group consisting of pFIFlac9, pFIFtrp69 and pFIFtrp^{3}69.

18. A microorganism transformed with an expression vehicle according to any one of claims 15-17.

19. The microorganism according to claim 18 obtained by transforming an E. coli strain.

20. A transformed microorganism according to claim 19 wherein said E. coli strain is E. coli K-12 strain 294.

21. A transformed microorganism according to claim 18 obtained by transforming Bacillus subtilis.

22. A transformed microorganism according to claim 18 obtained by transforming Saccharomyces cerevisiae.

***

Protein	1 2 3 4 Met-Ser-Tyr-Asn-	
mRNA	(5') AUG-UC$\begin{matrix}G\\A\\U\\C\end{matrix}$-UA$\begin{matrix}U\\C\end{matrix}$-AA$\begin{matrix}U\\C\end{matrix}$	(16 combinations)
	(5') AUG-AG$\begin{matrix}U\\C\end{matrix}$-UA$\begin{matrix}U\\C\end{matrix}$-AA$\begin{matrix}C\\U\end{matrix}$	(8 combinations)
Complementary DNA primers	ATT-$\begin{matrix}A\\G\end{matrix}$TA-$\begin{matrix}T\\C\end{matrix}$GA-CAT	Pool 1
	ATT-$\begin{matrix}A\\G\end{matrix}$TA-$\begin{matrix}A\\G\end{matrix}$GA-CAT	Pool 2
	ATT-$\begin{matrix}A\\G\end{matrix}$TA-$\begin{matrix}A\\G\end{matrix}$CT-CAT	Pool 3
	GTT-$\begin{matrix}A\\G\end{matrix}$TA-$\begin{matrix}T\\C\end{matrix}$GA-CAT	Pool 4
	GTT-$\begin{matrix}A\\G\end{matrix}$TA-$\begin{matrix}A\\G\end{matrix}$GA-CAT	Pool 5
	GTT-$\begin{matrix}A\\G\end{matrix}$TA-$\begin{matrix}A\\G\end{matrix}$CT-CAT	Pool 6

FIG. 1.

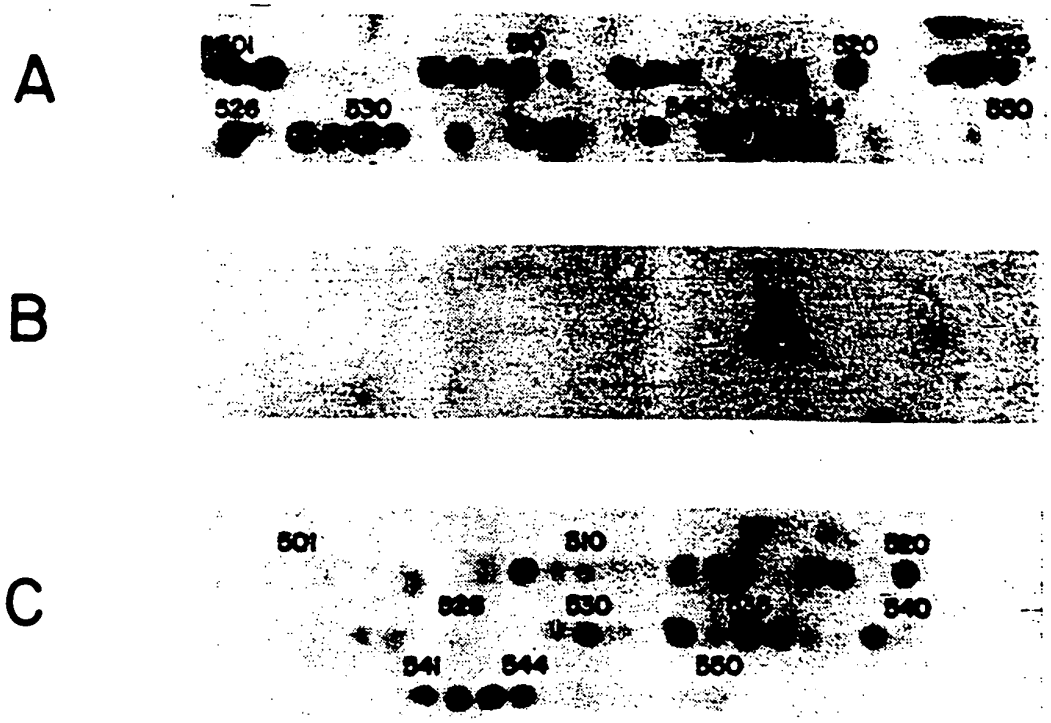

A
501
510
520
526
530
540
550
B
C
501
510
520
526
530
540
541
544
550

FIG. 2.

```
5'
S1                                        S10                                                   S20 S21
met thr asn lys cys leu leu gln ile ala leu leu leu cys phe ser thr thr ala leu ser MET SER TYR ASN
ATG ACC AAC AAG TGT CTC CTC CAA ATT GCT CTC CTG TTG TGC TTC TCC ACT ACA GCT CTT TCC ATG AGC TAC AAC
                                                                          50

                         10                                                  20
LEU LEU GLY PHE LEU GLN ARG SER SER ASN PHE GLN CYS GLN LYS LEU LEU TRP GLN LEU ASN GLY ARG LEU GLU
TTG CTT GGA TTC CTA CAA AGA AGC AGC AAT TTT CAG TGT CAG AAG CTC CTG TGG CAA TTG AAT GGG AGG CTT GAA
                              100                                                              150

30                                        40                                                  50
TYR CYS LEU LYS ASP ARG MET ASN PHE ASP ILE PRO GLU GLU ILE LYS GLN LEU GLN GLN PHE GLN LYS GLU ASP
TAT TGC CTC AAG GAC AGG ATG AAC TTT GAC ATC CCT GAG GAG ATT AAG CAG CTG CAG CAG TTC CAG AAG GAG GAC
                                                            200

                         60                                                  70
ALA ALA LEU THR ILE TYR GLU MET LEU GLN ASN ILE PHE ALA ILE PHE ARG GLN ASP SER SER SER THR GLY TRP
GCC GCA TTG ACC ATC TAT GAG ATG CTC CAG AAC ATC TTT GCT ATT TTC AGA CAA GAT TCA TCT AGC ACT GGC TGG
                              250                                                              300

80                                        90                                                  100
ASN GLU THR ILE VAL GLU ASN LEU LEU ALA ASN VAL TYR HIS GLN ILE ASN HIS LEU LYS THR VAL LEU GLU GLU
AAT GAG ACT ATT GTT GAG AAC CTC CTG GCT AAT GTC TAT CAT CAG ATA AAC CAT CTG AAG ACA GTC CTG GAA GAA
                                                      350

                         110                                                 120
LYS LEU GLU LYS GLU ASP PHE THR ARG GLY LYS LEU MET SER SER LEU HIS LEU LYS ARG TYR TYR GLY ARG ILE
AAA CTG GAG AAA GAA GAT TTT ACC AGG GGA AAA CTC ATG AGC AGT CTG CAC CTG AAA AGA TAT TAT GGG AGG ATT
                              400                                                              450

130                                       140                                                 150
LEU HIS TYR LEU LYS ALA LYS GLU TYR SER HIS CYS ALA TRP THR ILE VAL ARG VAL GLU ILE LEU ARG ASN PHE
CTG CAT TAC CTG AAG GCC AAG GAG TAC AGT CAC TGT GCC TGG ACC ATA GTC AGA GTG GAA ATC CTA AGG AAC TTT
                                                      500

                         160                 166
TYR PHE ILE ASN ARG LEU THR GLY TYR LEU ARG ASN END
TAC TTC ATT AAC AGA CTT ACA GGT TAC CTC CGA AAC TGA AGATCTCCTAGCCTGTCCCTCTGGGACTGGACAATTGCTTCAAGCA
                              550                                                              600

TTCTTCAACCAGCAGATGCTGTTTAAGTGACTGATGGCTAATGTACTGCAAATGAAAGGACACTAGAAGATTTTGAAATTTTTATTAAATTATGAGTT
                                        650                                                    700

ATTTTTATTTATTTAAATTTTATTTTGGAAAATAAATTATTTTTGGTGCAAAA
                                        750
                                                    3'
```

FIG. 3.

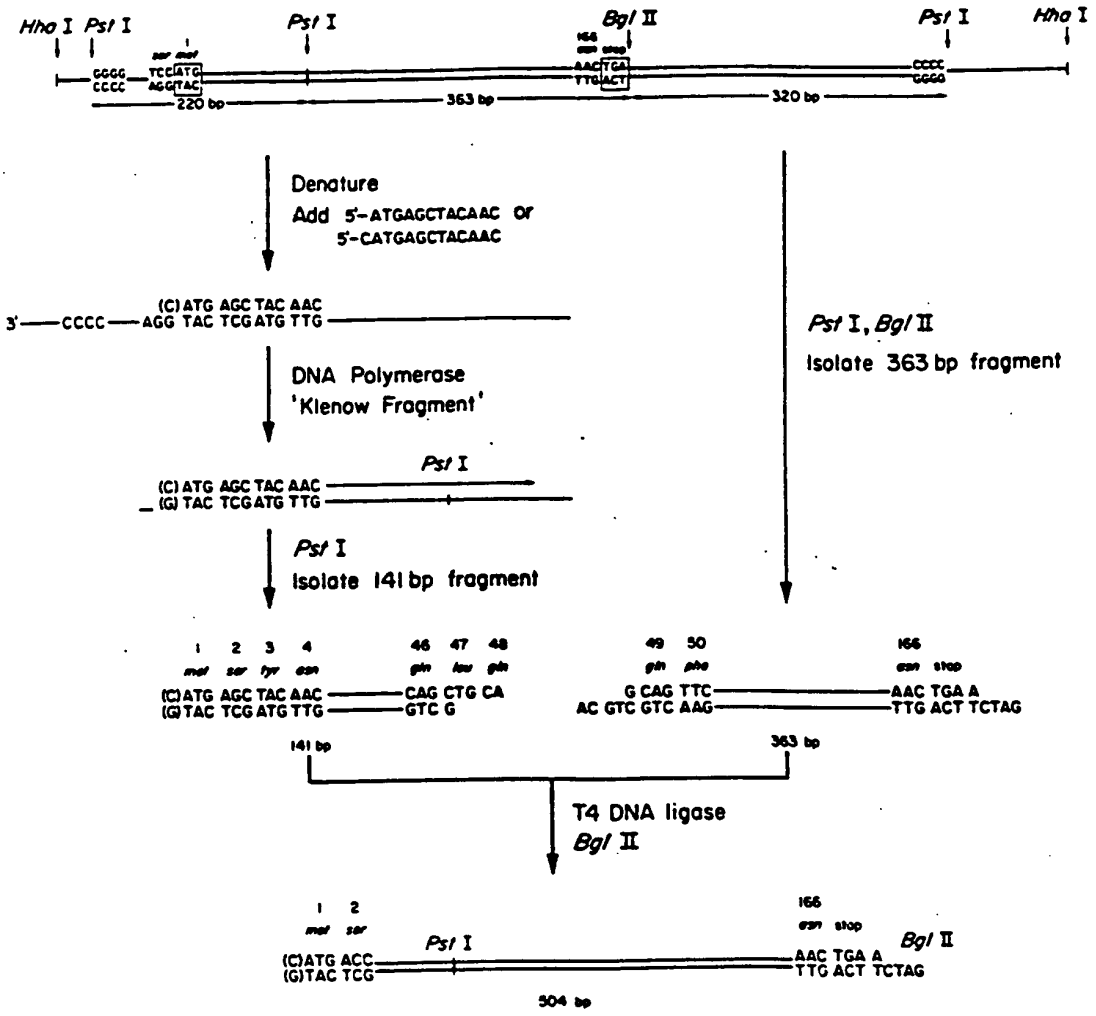

Hho I
Pst I
Pst I
Bgl II
Pst I
Hho I
met ser
ser stop
GGGG
CCCC
TCCATG
AGGTAC
AACTGA
TTGACT
CCCC
GGGG
220 bp
363 bp
320 bp
Denature
Add 5'-ATGAGCTACAAC or
5'-CATGAGCTACAAC
3'
CCCC
(C)ATG AGC TAC AAC
AGG TAC TCG ATG TTG
DNA Polymerase
'Klenow Fragment'
Pst I
(C)ATG AGC TAC AAC
(G)TAC TCG ATG TTG
Pst I
Isolate 141 bp fragment
Pst I, Bgl II
Isolate 363 bp fragment
1 2 3 4
met ser tyr asn
46 47 48
gln leu gln
(C)ATG AGC TAC AAC
(G)TAC TCG ATG TTG
CAG CTG CA
GTC G
49 50
gln phe
166
asn stop
G CAG TTC
AC GTC GTC AAG
AAC TGA A
TTG ACT TCTAG
141 bp
363 bp
T4 DNA ligase
Bgl II
1 2
met ser
166
asn stop
Pst I
Bgl II
(C)ATG ACC
(G)TAC TCG
AAC TGA A
TTG ACT TCTAG
504 bp

FIG. 4.

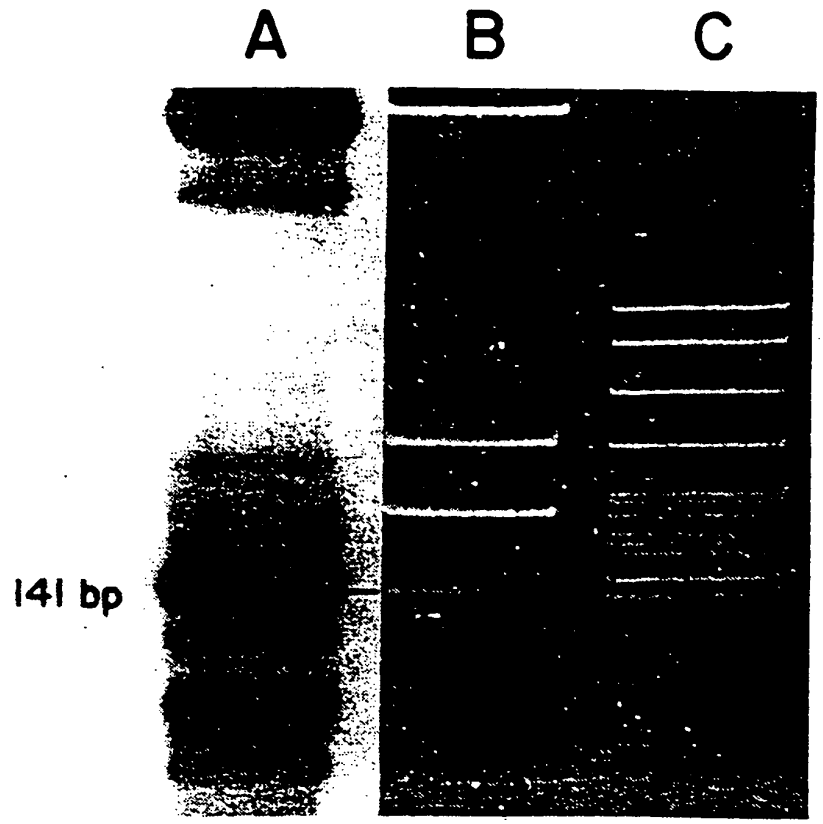
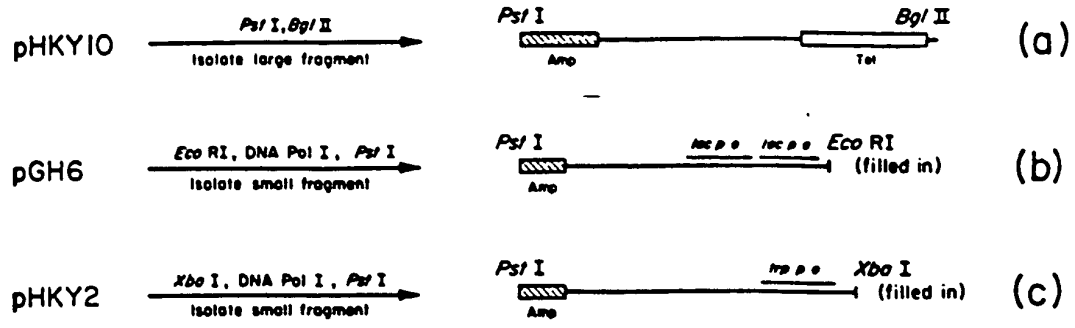


FIG. 5.

81.107621.5
— GENENTECH, Inc.


pHKY10
Pst I, Bgl II
Isolate large fragment
Pst I
Bgl II
Amp
Tet
(a)
pGH6
Eco RI, DNA Pol I, Pst I
Isolate small fragment
Pst I
lac p o
lac p o
Eco RI
(filled in)
Amp
(b)
pHKY2
Xba I, DNA Pol I, Pst I
Isolate small fragment
Pst I
trp p o
Xba I
(filled in)
Amp
(c)

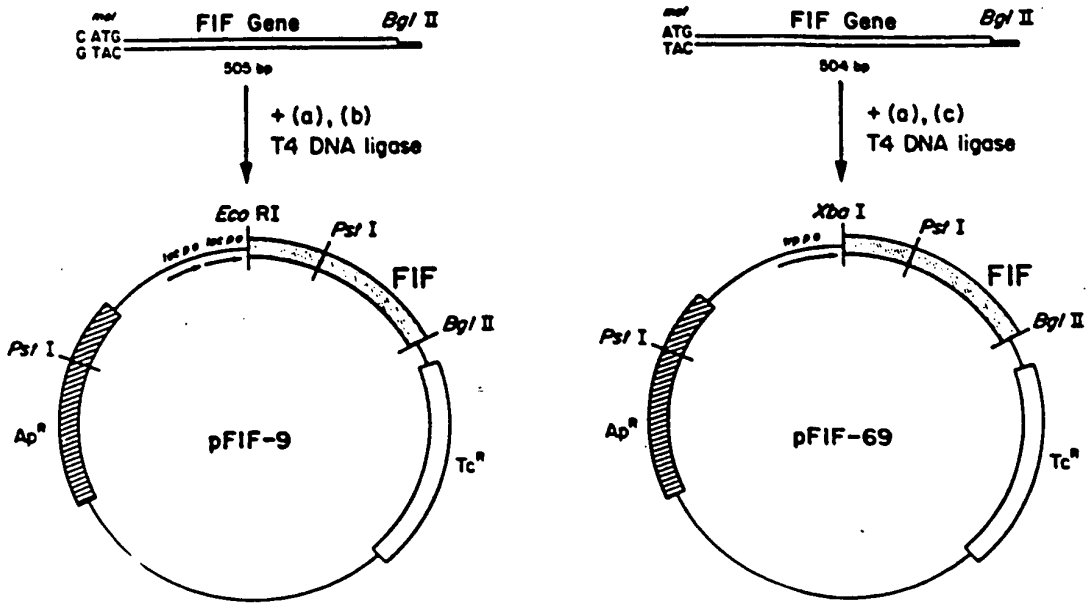

met
C ATG
G TAC
FIF Gene
Bgl II
505 bp
+ (a), (b)
T4 DNA ligase
Eco RI
Pst I
lac p o
lac p o
FIF
Bgl II
Pst I
Ap^R
pFIF-9
Tc^R
met
ATG
TAC
FIF Gene
Bgl II
504 bp
+ (a), (c)
T4 DNA ligase
Xba I
Pst I
trp p o
FIF
Bgl II
Pst I
Ap^R
pFIF-69
Tc^R

FIG. 6.